

High order splitting schemes with complex timesteps and their application in mathematical finance

Philipp Dörsek

Departement Mathematik, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland

Eskil Hansen

Centre for Mathematical Sciences, Lund University, P.O. Box 118, SE-22100 Lund, Sweden

Abstract

High order splitting schemes with complex timesteps are applied to Kolmogorov backward equations stemming from stochastic differential equations in Stratonovich form. In the setting of weighted spaces, the necessary analyticity of the split semigroups can be easily proved. A numerical example from interest rate theory, the CIR2 model, is considered. The numerical results are robust for drift-dominated problems, and confirm our theoretical results.

Keywords: splitting methods, complex coefficients, mathematical finance, convection-dominated problems, interest rate theory

1. Introduction

In mathematical finance, the pricing of derivative contracts can be reduced to the calculation of expected values under the risk-neutral measure, see, e.g., [3, 38]. This can be performed in different manners. The most general approach is to use Monte Carlo methods, see [27, 18], where the recently introduced multilevel Monte Carlo approach from [17] led to fast methods for problems with very low smoothness.

In contrast, we want to obtain approximations of the solution of the Kolmogorov backward PDE by splitting schemes. Such methods were also applied successfully in Quasi Monte Carlo simulations in [32], but all such approaches are in the end limited by the accuracy of the integration scheme. Furthermore, it is not straightforward to evaluate the stochastic processes at complex timesteps, which is necessary if splittings of order higher than 2 are to be used.

We therefore solve the PDE by finite element methods in space and a high order splitting method in time. Such high order splitting methods, overcoming the order barrier of 2 commonly encountered for splittings with nonnegative times, see [4], were introduced in [6, 23] and make use of the analyticity of the split semigroups. To show this analyticity, we make use of function spaces endowed with weighted supremum norms, originally introduced in [35] for proving

the existence of solutions to martingale problems for stochastic partial differential equations, and used for the analysis of numerical methods for stochastic partial differential equations in [12, 11, 10, 13, 14]. It turns out that using this framework, it is very simple to prove analyticity for semigroups stemming from stochastic differential equations in Stratonovich form, hence optimal rates of convergence follow. In particular, these results apply to problems on unbounded domains with unbounded coefficients vanishing at the (finite) boundaries of the domain. Such problems are usually difficult to deal with in Sobolev spaces and require the use of weighted Sobolev norms vanishing at the finite boundaries, see [8, 9] for some recent results on the Heston stochastic volatility model and a thorough discussion of references on this subject.

A particularly interesting feature of the considered numerical method is that the drift part is completely separated from the diffusion part. This means that we can choose suitable numerical schemes for each of these parts separately. In particular, if the first order hyperbolic drift part is solved by a method such as streamline diffusion finite elements or a discontinuous Galerkin method, we can expect that the method is robust for vanishing diffusivity. The need for such methods, notably for applications in mathematical finance, was recently observed in [30]. Other methods yielding such robustness are streamline diffusion methods, see [31, 16], and discontinuous Galerkin methods, see, e.g., [7, 24]. We stress that an advantage of our method is that different, optimised solvers for the drift part and the diffusion part can be used, providing a more flexible scheme.

The paper is organised as follows. Section 2 recalls the definitions of weighted spaces, specialising the results from [35] and [12, 13] to the finite-dimensional case. Furthermore, we show analyticity for a wide class of Markov semigroups in the setting of weighted spaces. Next, Section 3 formulates the splitting scheme and contains a convergence result. Finally, we show numerical findings for the CIR2 model in Section 4. In particular, we observe robustness of the suggested method for drift-dominated problems.

2. The functional analytic framework

Let us start off by recalling some basic facts from stochastic analysis, see, e.g., [26, 33, 34] for more details. Fix a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ and a d -dimensional standard Brownian motion $(W_t^j)_{j=1, \dots, d, t \geq 0}$ on it. Let $N \in \mathbb{N}$. Consider a stochastic differential equation in Stratonovich form on the closure $D \subset \mathbb{R}^N$ of a (not necessarily bounded) Lipschitz domain,

$$dX(t, x) = \sum_{j=0}^d V_j(X(t, x)) \circ dW_t^j, \quad X(0, x) = x, \quad (1)$$

where $W_t^0 = t$ and $\circ dW_t^0 = dt$, $V_j: D \rightarrow \mathbb{R}^N$ are Lipschitz continuous vector fields with appropriate smoothness assumptions discussed below in more detail, and $x \in D$. We suppose that the solution $(X(t, x))_{t \geq 0}$ is well-defined in D , in

particular that up to some common null set $N \in \mathcal{F}$, $\mathbb{P}(N) = 0$, $X(t, x) \in D$ for all $x \in D$ and $t \geq 0$. This is justified by well-known results on stochastic flows, see, e.g., [34, Section V.7]. For $f \in C^2(D)$ with suitable growth at infinity, Itô's lemma shows that $u(t, x) := \mathbb{E}[f(X(T-t, x))]$ satisfies the backward Kolmogorov equation

$$\frac{d}{dt}u(t, x) + \mathcal{L}u(t, x) = 0, \quad t > 0, \quad x \in D, \quad (2a)$$

$$u(T, x) = f(x), \quad x \in D. \quad (2b)$$

Here, \mathcal{L} denotes the “sum of squares” partial differential operator, defined for $g \in C^2(D)$ by

$$\mathcal{L}g(x) := V_0g(x) + \frac{1}{2} \sum_{j=1}^d V_j^2g(x), \quad (3)$$

with $Vg(x) := V(x) \cdot \nabla g(x)$ the directional derivative for $V: D \rightarrow \mathbb{R}^N$ and $g \in C^1(D)$. We split this operator into

$$\mathcal{L}_0g(x) := V_0g(x) \quad \text{and} \quad \mathcal{L}_1g(x) := \frac{1}{2} \sum_{j=1}^d V_j^2g(x). \quad (4)$$

The respective split stochastic differential equations read

$$\frac{d}{dt}X^0(t, x) = V_0(X^0(t, x)), \quad X^0(0, x) = x, \quad \text{and} \quad (5)$$

$$dX^1(t, x) = \sum_{j=1}^d V_j(X^1(t, x)) \circ dW_t^j, \quad X^1(0, x) = x, \quad (6)$$

cf. also [32].

We recall the following definitions from [35] and [12, 13, 14].

Definition 1. A function $\psi: D \rightarrow (0, \infty)$ is called *D-admissible weight function* if $\lim_{\|x\| \rightarrow \infty} \psi(x) = \infty$ and ψ is bounded on compact sets, where $\|\cdot\|$ denotes the Euclidean norm.

Note that the definition from [14] simplifies in this case, as D is locally compact. While the results in [35] and [12, 14] are stated for real-valued functions only, they also hold true for the complex-valued versions of the spaces considered here.

Definition 2. Fix $k \in \mathbb{N}_0$. For $j = 0, \dots, k$, let $\psi_j: D \rightarrow (0, \infty)$ be D-admissible weight functions. The space $\mathcal{B}_k^{(\psi_j)_{j=0, \dots, k}}(D)$ is defined as the closure of $C_b^k(D)$, the space of functions $f: D \rightarrow \mathbb{C}$ such that f is bounded and k times differentiable with all derivatives up to order k continuous and bounded, with respect to the norm $\|\cdot\|_{(\psi_j)_{j=0, \dots, k}, k}$, where

$$\|f\|_{(\psi_j)_{j=0, \dots, k}, k} := \|f\|_{\psi_0} + \sum_{j=1}^k |f|_{\psi_j, j} \quad (7)$$

with

$$\|f\|_{\psi_0} := \sup_{x \in D} \psi_0(x)^{-1} |f(x)| \quad \text{and} \quad \|f\|_{\psi_j, j} := \sup_{x \in D} \psi_j(x)^{-1} \|D^j f(x)\|. \quad (8)$$

Here, $D^j f(x)$ is seen as a j -linear form $(\mathbb{C}^N)^j \rightarrow \mathbb{C}$, endowed with the norm

$$\|D^j f(x)\| := \sup_{\substack{z_1, \dots, z_j \in \mathbb{C}^N \\ \|z_1\|, \dots, \|z_j\| \leq 1}} |D^j f(x)(z_1, \dots, z_j)|. \quad (9)$$

We also write $\mathcal{B}^\psi(D) := \mathcal{B}_0^\psi(D)$.

It is shown in [10, 14] that $\mathcal{B}_k^{(\psi_j)_{j=0, \dots, k}}(D) \subset C^k(D)$, and that it is a Banach space. Furthermore, the results of [35, 12, 10] yield the following generalised Feller property.

Proposition 3. *Let ψ be a D -admissible weight function. Suppose that $(P_t)_{t \geq 0}$ is a family of bounded linear operators on $\mathcal{B}_0^\psi(D)$ such that*

- $P_0 = I$, the identity on $\mathcal{B}^\psi(D)$,
- $P_{t+s} = P_t P_s$ for all $t, s \geq 0$, and
- $\lim_{t \rightarrow 0} P_t f(x) = f(x)$ for all $x \in D$ and $f \in \mathcal{B}^\psi(D)$.

Then, $(P_t)_{t \geq 0}$ is strongly continuous on $\mathcal{B}^\psi(D)$.

For simplicity, we formulate the following result for polynomially growing functions and vector fields with bounded derivatives. Therefore, we fix $s \in \mathbb{N}$ large enough and set $\psi_\ell(x) := (1 + \|x\|^2)^{(s-\ell)/2}$, $\ell \in \mathbb{N}_0$ and $\psi(x) := \psi_0(x)$. Similar results can be obtained for different choices, see, e.g., [13] for the case of exponentially growing functions and bounded vector fields. It is straightforward to see that $P_t f(x) := \mathbb{E}[f(X(t, x))]$, $P_t^0 f(x) := \mathbb{E}[f(X^0(t, x))]$ and $P_t^1 f(x) := \mathbb{E}[f(X^1(t, x))]$ define strongly continuous semigroups on $\mathcal{B}^\psi(D)$, and that their generators are suitable extensions of \mathcal{L} , \mathcal{L}_0 and \mathcal{L}_1 , respectively. Here, we apply the usual time inversion to turn the final value problem (2) into an initial value problem.

Proposition 4. *Let $k \in \mathbb{N}$. Assume that for $j = 0, \dots, d$, $V_j \in C^{k+1}(D; \mathbb{R}^N)$ with all derivatives of order 1 to $k+1$ bounded. Then, $P_t C_b^k(D) \subset C_b^k(D)$ for all $t \geq 0$, and*

$$\|P_t f\|_{(\psi_j)_{j=0, \dots, k}} \leq C(t) \|f\|_{(\psi_j)_{j=0, \dots, k}} \quad \text{for all } f \in \mathcal{B}_k^{(\psi_j)_{j=0, \dots, k}}(D), \quad (10)$$

where $C: [0, \infty) \rightarrow (0, \infty)$ is monotonously increasing. In particular,

$$P_t: \mathcal{B}_k^{(\psi_j)_{j=0, \dots, k}}(D) \rightarrow \mathcal{B}_k^{(\psi_j)_{j=0, \dots, k}}(D) \quad (11)$$

is a bounded operator with norm $\leq C(t)$, $t \geq 0$.

Proof. The first part and the norm bound is a consequence of [28, Theorem 3.3, p. 223], and the invariance of $\mathcal{B}_k^{(\psi_j)_{j=0,\dots,k}}(D)$ follows by a density argument. See also [13, Lemma 13] for a related result. \square

Proposition 5. *Let $V: D \rightarrow \mathbb{R}^N$ be a continuously differentiable vector field with bounded derivative. Suppose that $V(x) \cdot n(x) \leq 0$ for almost every $x \in \partial D$, where $n: \partial D \rightarrow \mathbb{R}^N$ is the (almost everywhere defined) outer unit normal vector to D . For $t \geq 0$, consider the semiflow $\text{Fl}_t^V: D \rightarrow D$ of V , i.e.,*

$$\frac{d}{dt}\text{Fl}_t^V(x) = V(\text{Fl}_t^V(x)), \quad \text{Fl}_0^V(x) = x. \quad (12)$$

Then,

$$P_t^V f(x) := f(\text{Fl}_t^V(x)), \quad t \geq 0, \quad (13)$$

defines a strongly continuous semigroup on $\mathcal{B}^\psi(D)$, and its generator is the closure of the operator $V: \mathcal{B}_1^{(\psi, \psi_1)}(D) \rightarrow \mathcal{B}^\psi(D)$. If $V(x) \cdot n(x) = 0$ on ∂D , $(P_t^V)_{t \geq 0}$ can be extended to a strongly continuous group.

Proof. The Lipschitz continuity of V together with $V(x) \cdot n(x) \leq 0$ for almost every $x \in \partial D$ yield that Fl_t^V is continuous and well-defined for $t \geq 0$, and that for some $\varepsilon > 0$ there exists $C > 0$ such that

$$\psi(\text{Fl}_t^V(x)) \leq C\psi(x) \quad \text{and} \quad (14)$$

$$\psi_1(\text{Fl}_t^V(x)) \leq C\psi_1(x) \quad \text{for all } t \in [0, \varepsilon] \text{ and } x \in D. \quad (15)$$

Hence, for arbitrary $f \in \mathcal{B}^\psi(D)$,

$$\sup_{x \in D} \psi(x)^{-1} |f(\text{Fl}_t^V(x))| \leq \sup_{x \in D} \psi(x)^{-1} \psi(\text{Fl}_t^V(x)) \|f\|_{\psi,0} \leq C \|f\|_{\psi,0}, \quad (16)$$

whence $P_t^V: \mathcal{B}^\psi(D) \rightarrow \mathcal{B}^\psi(D)$ is bounded for $t \geq 0$. As the semigroup property of $(P_t^V)_{t \in \mathbb{R}}$ is obvious, we can apply Proposition 3 to obtain that $(P_t^V)_{t \geq 0}$ indeed is a strongly continuous semigroup on $\mathcal{B}^\psi(D)$.

Let us now determine its generator \mathcal{G}^V . We shall prove that $\mathcal{B}_1^{(\psi, \psi_1)}(D) \subset \text{dom } \mathcal{G}^V$, and that $\mathcal{G}^V f(x) = Vf(x)$ on this space. Clearly, for $f \in C_b^1(D)$, $\frac{d}{dt}|_{t=0} f(\text{Fl}_t^V(x)) = Vf(x)$, and $Vf \in \mathcal{B}^\psi(D)$. Hence,

$$P_t^V f - f = \int_0^t P_s^V (Vf) ds, \quad (17)$$

and this implies that $f \in \text{dom } \mathcal{G}^V$ and $\mathcal{G}^V f = Vf$. The Lipschitz continuity of V implies $\|V(x)\| \leq C(1 + \|x\|^2)^{1/2}$ with some constant $C > 0$, and we observe

$$|Vf(x)| \leq C(1 + \|x\|^2)^{1/2} \psi_1(x) |f|_{\psi_1,1} \leq C\psi(x) |f|_{\psi_1,1}. \quad (18)$$

It follows that $V: \mathcal{B}_1^{(\psi, \psi_1)}(D) \rightarrow \mathcal{B}^\psi(D)$ is bounded, and together with the closedness of \mathcal{G}^V and Proposition 11 this implies $\mathcal{B}_1^{(\psi, \psi_1)}(D) \subset \text{dom } \mathcal{G}^V$ and

$$\mathcal{G}^V f = Vf \quad \text{for all } f \in \mathcal{B}_1^{(\psi, \psi_1)}(D). \quad (19)$$

If $f \in \mathcal{B}_1^{(\psi, \psi_1)}(D)$, we see that $x \mapsto P_t^V f(x)$ is continuously differentiable, and that for $z \in \mathbb{C}^N$,

$$|DP_t^V f(x)z| = |Df(\text{Fl}_t^V(x))D\text{Fl}_t^V(x)z| \leq |f|_{\psi_1, 1} \psi_1(\text{Fl}_t^V(x)) |D\text{Fl}_t^V(x)z|. \quad (20)$$

Denote the Lipschitz constant of V by L , then by Gronwall's inequality,

$$|D\text{Fl}_t^V(x)z| \leq \exp(L|t|) \|z\|. \quad (21)$$

Hence, for $t \in [0, \varepsilon]$, $\|P_t^V f\|_{(\psi, \psi_1), 1} \leq C \|f\|_{(\psi, \psi_1), 1}$. As additionally $P_t^V(C_b^1(D)) \subset C_b^1(D)$, we obtain that $\mathcal{B}_1^{(\psi, \psi_1)}(D)$ is a core for \mathcal{G}^V by [15, Proposition II.1.7].

If we assume that $V(x) \cdot n(x) = 0$ for $x \in \partial D$, we can apply the above results to both V and $-V$, from which it follows that $(P_t^V)_{t \in \mathbb{R}}$ is a strongly continuous group. The proof is thus complete. \square

Corollary 6. *Let $V \in C^3(D; \mathbb{R}^N)$ with bounded derivatives of order 1 to 3 such that $V(x) \cdot n(x) = 0$ for almost every $x \in \partial D$. Then,*

$$P_t^{\frac{1}{2}V^2} f(x) := \int_{\mathbb{R}} f(\text{Fl}_{\sqrt{t}y}^V(x)) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy \quad (22)$$

defines an analytic semigroup of angle $\pi/2$ on $\mathcal{B}^\psi(D)$. Its generator is the closure of the operator $\frac{1}{2}V^2: \mathcal{B}_2^{(\psi, \psi_1, \psi_2)}(D) \rightarrow \mathcal{B}^\psi(D)$.

Proof. This is a consequence of Proposition 5 and [15, Corollary II.4.9]. That $\mathcal{B}_2^{(\psi, \psi_1, \psi_2)}(D)$ is a core is a consequence of Proposition 4. \square

Corollary 7. *Suppose that the vector field V_0 is continuously differentiable with bounded derivative and that the vector fields V_j , $j = 1, \dots, d$ are three times continuously differentiable with bounded derivatives of order 1 to 3 on the Lipschitz domain $D \subset \mathbb{R}^N$, and satisfy*

$$V_0(x) \cdot n(x) \leq 0 \quad \text{and} \quad V_j(x) \cdot n(x) = 0 \quad \text{for } j = 1, \dots, d \text{ and } x \in \partial D, \quad (23)$$

where $n(x)$ denotes the (almost everywhere defined) outer unit normal vector on D in $x \in \partial D$. Define $\mathcal{L}_0 g(x) := V_0 g(x)$ and $\mathcal{L}_1 g(x) := \frac{1}{2} \sum_{j=1}^d V_j^2 g(x)$ for $g \in \mathcal{B}_2^{(\psi_j)_{j=0,1,2}}(D)$. Then, the closure of \mathcal{L}_0 generates the strongly continuous semigroup $(P_t^0)_{t \geq 0}$, and the closure of \mathcal{L}_1 generates the analytical semigroup of angle $\pi/2$ $(P_t^1)_{t \geq 0}$ on $\mathcal{B}^\psi(D)$.

Proof. This follows from a combination of Proposition 5, Corollary 6, and Proposition 12. \square

Remark 8. We only obtained properties of the split semigroups $(P_t^0)_{t \geq 0}$ and $(P_t^1)_{t \geq 0}$ in Corollary 7. While perturbation theory arguments should allow us to derive properties of the semigroup $(P_t)_{t \geq 0}$ from these results by applying perturbation theory, see, e.g., [15, Theorem III.2.10], this question is not central for the topic of this paper and is therefore left to further research.

3. High order splitting schemes

Recently, it was observed in [23, 6] that complex timesteps can be used to obtain splitting schemes with order higher than 2 for parabolic problems. In the given setting, we have the following result.

Theorem 9. *Let $(\alpha_k)_{k=1,\dots,s} \subset [0, \infty)$, $(\beta_k)_{k=1,\dots,s} \subset \{z \in \mathbb{C}: \operatorname{Re} z \geq 0\}$ be the coefficients of a splitting scheme of formal order p . Assume the setting of Corollary 7, and that additionally, V_j for $j = 0, \dots, d$ are $2p + 3$ times continuously differentiable with all derivatives of order 1 to $2p + 3$ bounded. For $f \in \mathcal{B}^\psi(D)$, set*

$$Q_{(\Delta t)}f := P_{\alpha_1 \Delta t}^0 P_{\beta_1 \Delta t}^1 \dots P_{\alpha_s \Delta t}^0 P_{\beta_s \Delta t}^1 f. \quad (24)$$

Then, for $f \in \mathcal{B}_{2(p+1)}^{(\psi_j)_{j=0,\dots,2(p+1)}}(D)$ and $T \geq 0$, there exists a constant $C_{T,f} > 0$ independent of $n \in \mathbb{N}$ such that

$$\|Q_{(T/n)}^n f - P_T f\|_\psi \leq C_{T,f} n^{-p}. \quad (25)$$

Proof. By Proposition 4, $\mathcal{B}_{2(p+1)}^{(\psi, \psi_1, \dots, \psi_{2(p+1)})}(D)$ is invariant with respect to the semigroup $(P_t)_{t \geq 0}$. Using the results from Section 2, this is hence a straightforward consequence of [22, Theorem 2.3], see also [23, Theorem 3.1] and [6, Theorem 3.1]. \square

Remark 10. A fourth order splitting as required by the above theorem was constructed in [6, equation (5.1)]. They have $s = 5$ and

$$\alpha_1 = 0, \quad \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \frac{1}{4}, \quad (26)$$

$$\beta_1 = \beta_5 = \frac{1}{10} - \frac{1}{30}i, \quad \beta_2 = \beta_4 = \frac{4}{15} + \frac{2}{15}i, \quad \text{and} \quad \beta_3 = \frac{4}{15} - \frac{1}{5}i. \quad (27)$$

This scheme requires the evaluation of the semigroups $P_{\alpha_k \Delta t}^0$ and $P_{\beta_k \Delta t}^1$ to fourth order. The strategy of composition methods as applied in [6] will fail if \mathcal{L}_0 does not generate a group, which is a consequence of [4]. As V_0 usually does not satisfy $V_0(x) \cdot n(x) = 0$ for $x \in \partial D$ in problems from mathematical finance with mean reversion because V_0 points towards the mean and hence \mathcal{L}_0 does not generate a group, we do not analyse this question further here.

4. Numerical example: The CIR2 model

The CIR2 (Longstaff–Schwartz) model, see [29, 5], is a popular model for interest rates. It supposes that the short rate r_t is given by

$$r_t = x_t + y_t \quad (28)$$

with x_t and y_t being square-root diffusions (CIR processes) under the risk-neutral measure \mathbb{Q} ,

$$dx_t = \theta_x(\mu_x - x_t)dt + \sigma_x \sqrt{x_t} dW_t^x, \quad (29a)$$

$$dy_t = \theta_y(\mu_y - y_t)dt + \sigma_y \sqrt{y_t} dW_t^y \quad (29b)$$

where the stochastic basis is given by $(\Omega, \mathcal{F}, \mathbb{Q}, (\mathcal{F}_t)_{t \geq 0})$, $\theta_x, \mu_x, \sigma_x, \theta_y, \mu_y$ and σ_y are positive constants, and W_t^x and W_t^y are independent Brownian motions. From $(r_t)_{t \geq 0}$, prices of default-free zero coupon bonds can be derived via

$$B(S, T) := \mathbb{E}[\exp(-\int_S^T r_t dt) | \mathcal{F}_S], \quad (30)$$

where $\mathbb{E}[\cdot | \mathcal{F}_S]$ denotes conditional expectation with respect to \mathcal{F}_S under \mathbb{Q} , and $0 \leq S \leq T$.

For simplicity, we consider the approximation of the bond price itself, for which analytic formulas are available. This allows us to determine the precise error of our numerical results. As the bond price has smooth dependence on the initial values x_0 and y_0 , we do not need to use non-uniform timesteps in the approximation. In the approximation of non-smooth payoffs, however, this is necessary; for a construction of non-uniform time grids, see, e.g., [20, 36, 37].

After the usual time inversion, the PDE formulation of the problem reads

$$\frac{\partial}{\partial t} u(t, x, y) = \mathcal{L}u(t, x, y), \quad u(0, x, y) = 1, \quad (31)$$

with $u(T, x, y) = B(0, T)$ if we fix $x_0 = x$ and $y_0 = y$, where

$$\begin{aligned} \mathcal{L}u(t, x, y) = & \theta_x(\mu_x - x) \frac{\partial}{\partial x} u(t, x, y) + \frac{1}{2} \sigma_x^2 x \frac{\partial^2}{\partial x^2} u(t, x, y) \\ & + \theta_y(\mu_y - y) \frac{\partial}{\partial y} u(t, x, y) + \frac{1}{2} \sigma_y^2 y \frac{\partial^2}{\partial y^2} u(t, x, y) - (x + y)u(t, x, y). \end{aligned} \quad (32)$$

We split this equation according to

$$\begin{aligned} \mathcal{L}_0 u_0(t, x, y) = & \left[\theta_x(\mu_x - x) - \frac{1}{4} \sigma_x^2 \right] \frac{\partial}{\partial x} u_0(t, x, y) \\ & + \left[\theta_y(\mu_y - y) - \frac{1}{4} \sigma_y^2 \right] \frac{\partial}{\partial y} u_0(t, x, y) \quad \text{and} \end{aligned} \quad (33)$$

$$\begin{aligned} \mathcal{L}_1 u_1(t, x, y) = & \frac{1}{2} \sigma_x^2 x \frac{\partial^2}{\partial x^2} u_1(t, x, y) + \frac{1}{4} \sigma_x^2 \frac{\partial}{\partial x} u_1(t, x, y) \\ & + \frac{1}{2} \sigma_y^2 y \frac{\partial^2}{\partial y^2} u_1(t, x, y) + \frac{1}{4} \sigma_y^2 \frac{\partial}{\partial y} u_1(t, x, y) - (x + y)u_1(t, x, y), \end{aligned} \quad (34)$$

which corresponds to a split into drift \mathcal{L}_0 and diffusion \mathcal{L}_1 after transformation into Stratonovich form. In order to ensure that $x = 0$ and $y = 0$ are outflow boundaries for the hyperbolic problem $\frac{\partial}{\partial t} u(t, x, y) = \mathcal{L}_0 u(t, x, y)$, we assume $4\theta_x \mu_x \geq \sigma_x^2$ and $4\theta_y \mu_y \geq \sigma_y^2$, which is weaker than the Feller condition, see, e.g., [2]. While the given operator does not satisfy the assumptions of Section 2 to prove that \mathcal{L}_1 generates an analytic semigroup, this can be proved directly. Proposition 13 shows this result for the 1D case, and the proof generalises to the

considered equation. The necessary smoothness of the exact solution $u(t, x, y)$ of equation (31) can also not be obtained directly from Proposition 4, but follows from [1, Proposition 43].

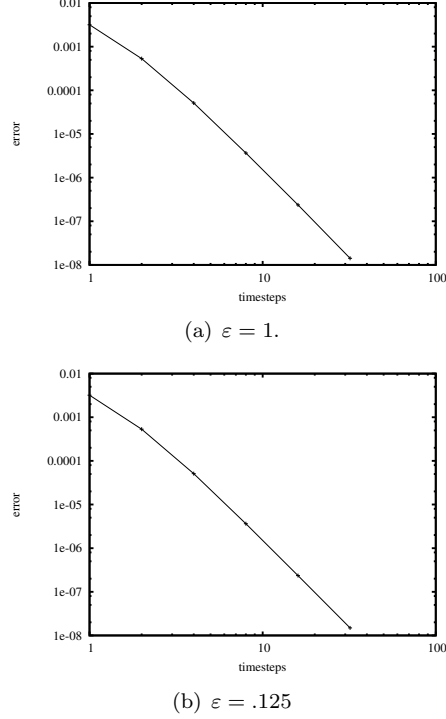


Figure 1: Errors in approximation of bond prices of the CIR2 model

In the computations, we fix the parameters

$$\theta_x = 15.5, \mu_x = .025, \sigma_x = .2\varepsilon, \theta_y = 20.5, \mu_y = .025, \sigma_y = .3\varepsilon, \quad (35)$$

where we consider $\varepsilon = 1.$ and $\varepsilon = .125.$ In both cases, the problem is drift-dominated. As we want to focus on the time discretisation error, we minimise the impact of the domain truncation by cutting off far beyond any reasonable values for the state variables, at $x = 16$ and $y = 16$, and of the spatial discretisation by using higher order continuous finite elements in space, with 17 grid points per direction and a polynomial degree of 4. This leads to a total of 4225 spatial degrees of freedom. The operator splitting uses the scheme from [6, equation (5.1)], see Remark 10 for the coefficients. For the split equations, we employ a streamline diffusion FEM [25] using a polynomial degree of 4 for the first order hyperbolic equation involving \mathcal{L}_0 , and a resolvent Krylov subspace method using a 30-dimensional Krylov subspace for the approximation of the matrix exponential [21, 19] of the second order parabolic equation involving \mathcal{L}_1 .

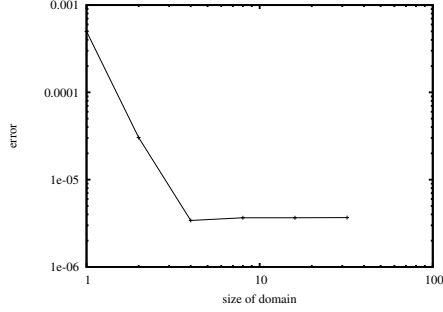


Figure 2: Effects of domain truncation

Figure 1 shows the results of a numerical computation. The plotted error is the difference between $B(0, T)$ and the numerical approximation $B_{\text{app}}(0, T)$, measured in the \mathcal{B}^ψ -norm with $\psi(x, y) = (1 + x^2 + y^2)^3$. We clearly observe 4th order convergence. If we restrict $x + y \leq 1$, which should include all values of practical interest, we see that $\psi(x, y) \leq 64$. Hence, in this region, the form of the error norm ensures that the pointwise approximation error is less than 10^{-3} using only 8 timesteps. Furthermore, the error is virtually independent of the value of ε , which means that our approximation is robust for the case of vanishing diffusion.

In order to understand the effect of the space truncation, we plot in Figure 2 the errors for different choices of the cutoff point, where the spatial mesh width is kept constant and 8 timesteps are used. We observe that if the domain size is increased, the error stays bounded.

Appendix A. Functional-analytic results

In this section, we collect a result on closed operators and one on analytic semigroups that are easy to prove, but that were not able to find in standard literature.

Proposition 11. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces such that Y is continuously embedded in X , i.e., $Y \subset X$ and $\|y\|_X \leq C\|y\|_Y$ for all $y \in Y$ with some constant $C > 0$, and Z be dense in Y (with respect to the norm of Y). Given a closed operator $A: \text{dom } A \subset X \rightarrow X$ with $Z \subset \text{dom } A$ such that $\|Az\|_X \leq K\|z\|_Y$ for all $z \in Z$ with $K \geq 0$, we have $Y \subset \text{dom } A$ and $\|Ay\|_X \leq K\|y\|_Y$ for all $y \in Y$.*

Proof. By continuity, $A|_Z$ can be extended to a continuous linear operator $B: Y \rightarrow X$ with $\|By\|_X \leq K\|y\|_Y$ and $Bz = Az$ for all $z \in Z$. Given $y \in Y$, let $(z_n)_{n \in \mathbb{N}}$ be such that $\lim_{n \rightarrow \infty} \|y - z_n\|_Y = 0$. Then, by continuity of B , $\lim_{n \rightarrow \infty} \|By - Bz_n\|_X = 0$. As $Bz_n = Az_n$ for all $n \in \mathbb{N}$, we observe $z_n \rightarrow y$ and $Az_n \rightarrow By$ in the norm of X . Hence, the closedness of A implies $y \in \text{dom } A$ and $Ay = By$, which yields the claim. \square

Proposition 12. Fix a complex Banach space X . Given $\delta \in (0, \pi/2]$, set $\Sigma_\delta := \{z \in \mathbb{C} : |\arg z| < \delta\}$. Let $(A, \text{dom } A)$, $(B, \text{dom } B)$ be the generators of analytic semigroups $(S_z^A)_{z \in \Sigma_\delta \cup 0}$ and $(S_z^B)_{z \in \Sigma_\delta \cup 0}$ on X satisfying

$$\|S_z^A\| \leq \exp(K|z|) \quad \text{and} \quad \|S_z^B\| \leq \exp(K|z|) \quad \text{for } z \in \Sigma_\delta \cup 0 \quad (\text{A.1})$$

for some $K > 0$, and assume that the closure $(C, \text{dom } C)$ of $(A + B, \text{dom } A \cap \text{dom } B)$ generates a strongly continuous semigroup $(S_t^C)_{t \geq 0}$.

Then, $(S_t^C)_{t \geq 0}$ extends to an analytic semigroup on $\Sigma_\delta \cup 0$.

Proof. We assume without loss of generality that the semigroups generated by A and B are bounded on Σ_δ ; otherwise, translate the operators appropriately first. By [15, Theorem II.4.6], we have to prove that for some $\theta \in (0, \pi/2)$, $e^{\pm i\theta}C$ generate strongly continuous semigroups. The Hille–Yosida theorem implies existence of $\omega_0 \in \mathbb{R}$ such that $\lambda - C$ is surjective for $\text{Re } \lambda \geq \omega_0$. As $\text{Re } e^{\pm i\theta} = \cos(\theta) > 0$, we can find $\mu > 0$ such that $\text{Re } \mu e^{\pm i\theta} \geq \omega_0$, whence $\mu - e^{\pm i\theta}C = e^{\pm i\theta}(e^{\mp i\theta}\mu - C)$ is invertible. Fixing $\theta \in (0, \delta)$, we can apply the Trotter product formula [15, Corollary III.5.8] to $e^{\pm i\theta}A$ and $e^{\pm i\theta}B$ to prove that $e^{\pm i\theta}C$ generate strongly continuous semigroups. Repeating this argument with A , B and C replaced by $e^{i\alpha}A$, $e^{i\alpha}B$ and $e^{i\alpha}C$, where $\alpha \in (-\delta, \delta)$, we obtain the result. \square

Appendix B. Analyticity of the CIR semigroup in the \mathcal{B}^ψ setting

The following result proves the analyticity of the CIR semigroup on a \mathcal{B}^ψ -space directly, as Corollary 6 does not apply due to the lacking Lipschitz continuity of the square root.

Proposition 13. Set $\psi(x) := (1 + x^2)^{s/2}$ with $s \in \mathbb{N}$. For $f \in \mathcal{B}^\psi([0, \infty))$, define

$$P_t f(x) := \int_{\mathbb{R}} f((\sqrt{x} + \sigma\sqrt{t}y)^2) \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy, \quad x \in [0, \infty), \quad (\text{B.1})$$

Then, $P_t f$ defines an analytic semigroup on $\mathcal{B}^\psi([0, \infty))$ and solves the PDE

$$\frac{\partial}{\partial t} u(t, x) = 2\sigma^2 x \frac{\partial^2}{\partial x^2} u(t, x) + \sigma^2 \frac{\partial}{\partial x} u(t, x), \quad t > 0 \text{ and } x \geq 0, \quad (\text{B.2a})$$

$$u(0, x) = f(x), \quad x \geq 0. \quad (\text{B.2b})$$

Proof. It is easy to see that $(P_t)_{t \geq 0}$ defines a generalised Feller semigroup on $\mathcal{B}^\psi([0, \infty))$, hence it is strongly continuous by Proposition 3. Integration by parts proves

$$\begin{aligned} \frac{\partial}{\partial t} P_t f(x) &= \int_{\mathbb{R}} f'((\sqrt{x} + \sigma\sqrt{t}y)^2) \frac{\sqrt{x} + \sigma\sqrt{t}y}{\sqrt{t}} \sigma y \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy \\ &= \frac{1}{2t} \int_{\mathbb{R}} y \frac{\partial}{\partial y} f((\sqrt{x} + \sigma\sqrt{t}y)^2) \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy \\ &= \frac{1}{2t} \int_{\mathbb{R}} (y^2 - 1) f((\sqrt{x} + \sigma\sqrt{t}y)^2) \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy, \end{aligned}$$

and as the right hand side defines a bounded operator on $\mathcal{B}^\psi([0, \infty))$ with norm of the order $O(t^{-1})$, $(P_t)_{t \geq 0}$ is an analytic semigroup.

To show that $P_t f$ satisfies the given PDE, we calculate

$$\begin{aligned} \frac{\partial}{\partial x} P_t f(x) &= \int_{\mathbb{R}} f'((\sqrt{x} + \sigma\sqrt{t}y)^2) \frac{\sqrt{x} + \sigma\sqrt{t}y}{\sqrt{x}} \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy \\ &= \frac{1}{2\sigma\sqrt{tx}} \int_{\mathbb{R}} \frac{\partial}{\partial y} f((\sqrt{x} + \sigma\sqrt{t}y)^2) \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy \\ &= \frac{1}{2\sigma\sqrt{tx}} \int_{\mathbb{R}} y f((\sqrt{x} + \sigma\sqrt{t}y)^2) \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy \end{aligned}$$

and, similarly,

$$\begin{aligned} \frac{\partial^2}{\partial x^2} P_t f(x) &= \frac{\partial}{\partial x} \left(\frac{1}{2\sigma\sqrt{tx}} \int_{\mathbb{R}} y f((\sqrt{x} + \sigma\sqrt{t}y)^2) \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy \right) \\ &= -\frac{1}{4\sigma\sqrt{tx^3}} \int_{\mathbb{R}} y f((\sqrt{x} + \sigma\sqrt{t}y)^2) \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy \\ &\quad + \frac{1}{4\sigma^2 tx} \int_{\mathbb{R}} y \frac{\partial}{\partial y} f((\sqrt{x} + \sigma\sqrt{t}y)^2) \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy \\ &= -\frac{1}{4\sigma\sqrt{tx^3}} \int_{\mathbb{R}} y f((\sqrt{x} + \sigma\sqrt{t}y)^2) \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy \\ &\quad + \frac{1}{4\sigma^2 tx} \int_{\mathbb{R}} (y^2 - 1) f((\sqrt{x} + \sigma\sqrt{t}y)^2) \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy. \end{aligned}$$

Hence,

$$\begin{aligned} \left(2\sigma^2 x \frac{\partial^2}{\partial x^2} + \sigma^2 \frac{\partial}{\partial x} \right) P_t f(x) &= \frac{1}{2t} \int_{\mathbb{R}} (y^2 - 1) f((\sqrt{x} + \sigma\sqrt{t}y)^2) \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy \\ &= \frac{\partial}{\partial t} P_t f(x), \end{aligned}$$

which proves the claim. \square

Acknowledgements

The first author gratefully acknowledges support by the ETH Foundation. The work of the second author was supported by the Swedish Research Council under grant 621-2011-5588.

References

- [1] Abdelkoddousse Ahdida and Aurélien Alfonsi. Exact and high order discretization schemes for Wishart processes and their affine extensions. *ArXiv e-prints*, June 2010.

- [2] Leif B.G. Andersen, Peter Jäckel, and Christian Kahl. *Simulation of Square-Root Processes*. John Wiley & Sons, Ltd, 2010.
- [3] Björk, Tomas. *Arbitrage theory in continuous time. 2nd ed.* Oxford: Oxford University Press, 2004.
- [4] Sergio Blanes and Fernando Casas. On the necessity of negative coefficients for operator splitting schemes of order higher than two. *Appl. Numer. Math.*, 54(1):23–37, 2005.
- [5] Damiano Brigo and Fabio Mercurio. *Interest rate models—theory and practice*. Springer Finance. Springer-Verlag, Berlin, 2001.
- [6] François Castella, Philippe Chartier, Stéphane Descombes, and Gilles Vilmart. Splitting methods with complex times for parabolic equations. *BIT*, 49(3):487–508, 2009.
- [7] Bernardo Cockburn and Chi-Wang Shu. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, 35(6):2440–2463 (electronic), 1998.
- [8] Panagiota Daskalopoulos and Paul M. N. Feehan. Existence, uniqueness, and global regularity for degenerate elliptic obstacle problems in mathematical finance. *ArXiv e-prints*, September 2011.
- [9] Panagiota Daskalopoulos and Paul M. N. Feehan. $C^{1,1}$ regularity for degenerate elliptic obstacle problems in mathematical finance. *ArXiv e-prints*, June 2012.
- [10] Philipp Dörsek. *Numerical Methods for Stochastic Partial Differential Equations*. PhD thesis, Vienna University of Technology, October 2011.
- [11] Philipp Dörsek. Semigroup Splitting and Cubature Approximations for the Stochastic Navier–Stokes Equations. *SIAM Journal on Numerical Analysis*, 50(2):729–746, 2012.
- [12] Philipp Dörsek and Josef Teichmann. A Semigroup Point Of View On Splitting Schemes For Stochastic (Partial) Differential Equations. *ArXiv e-prints*, November 2010.
- [13] Philipp Dörsek and Josef Teichmann. Efficient simulation and calibration of general HJM models by splitting schemes. *ArXiv e-prints*, December 2011.
- [14] Philipp Dörsek, Josef Teichmann, and Dejan Velušček. Cubature Methods For Stochastic (Partial) Differential Equations In Weighted Spaces. *ArXiv e-prints*, January 2012.

- [15] Klaus-Jochen Engel and Rainer Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.
- [16] Klaus Gerdes, Jens Markus Melenk, Christoph Schwab, and Dominik Schötzau. The hp -version of the streamline diffusion finite element method in two space dimensions. *Math. Models Methods Appl. Sci.*, 11(2):301–337, 2001.
- [17] Michael B. Giles. Multilevel Monte Carlo path simulation. *Oper. Res.*, 56(3):607–617, 2008.
- [18] Paul Glasserman. *Monte Carlo methods in financial engineering*. New York, NY: Springer, 2004.
- [19] Volker Grimm. Resolvent krylov subspace approximation to operator functions. *BIT Numerical Mathematics*, pages 1–21. 10.1007/s10543-011-0367-8.
- [20] Wen Zhuang Gui and Ivo Babuška. The h , p and h - p versions of the finite element method in 1 dimension. II. The error analysis of the h - and h - p versions. *Numer. Math.*, 49(6):613–657, 1986.
- [21] Stefan Güttel. *Rational Krylov methods for operator functions*. PhD thesis, TU Bergakademie Freiberg, 2010.
- [22] Eskil Hansen and Alexander Ostermann. Exponential splitting for unbounded operators. *Math. Comp.*, 78(267):1485–1496, 2009.
- [23] Eskil Hansen and Alexander Ostermann. High order splitting methods for analytic semigroups exist. *BIT*, 49(3):527–542, 2009.
- [24] Jan S. Hesthaven and Tim Warburton. *Nodal discontinuous Galerkin methods*, volume 54 of *Texts in Applied Mathematics*. Springer, New York, 2008. Algorithms, analysis, and applications.
- [25] Paul Houston, Christoph Schwab, and Endre Süli. Stabilized hp -finite element methods for first-order hyperbolic problems. *SIAM J. Numer. Anal.*, 37(5):1618–1643 (electronic), 2000.
- [26] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.
- [27] Peter E. Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.

- [28] Hiroshi Kunita. Stochastic differential equations and stochastic flows of diffeomorphisms. In *École d'été de probabilités de Saint-Flour, XII—1982*, volume 1097 of *Lecture Notes in Math.*, pages 143–303. Springer, Berlin, 1984.
- [29] Francis A. Longstaff and Eduardo S. Schwartz. Interest rate volatility and the term structure: A two-factor general equilibrium model. *The Journal of Finance*, 47(4):pp. 1259–1282, 1992.
- [30] Daniele Marazzina, Oleg Reichmann, and Christoph Schwab. *hp*-DGFEM for Kolmogorov-Fokker-Planck equations of multivariate Lévy processes. *Math. Models Methods Appl. Sci.*, 22(1):1150005, 37, 2012.
- [31] J. M. Melenk and C. Schwab. The *hp* streamline diffusion finite element method for convection dominated problems in one space dimension. *East-West J. Numer. Math.*, 7(1):31–60, 1999.
- [32] Syoiti Ninomiya and Nicolas Victoir. Weak approximation of stochastic differential equations and application to derivative pricing. *Appl. Math. Finance*, 15(1-2):107–121, 2008.
- [33] Bernt Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.
- [34] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2004. Stochastic Modelling and Applied Probability.
- [35] Michael Röckner and Zeev Sobol. Kolmogorov equations in infinite dimensions: well-posedness and regularity of solutions, with applications to stochastic generalized Burgers equations. *Ann. Probab.*, 34(2):663–727, 2006.
- [36] Dominik Schötzau. *hp-DGFEM for parabolic evolution problems*. PhD thesis, ETH Zürich, 1999.
- [37] Dominik Schötzau and Christoph Schwab. Time discretization of parabolic problems by the *hp*-version of the discontinuous Galerkin finite element method. *SIAM J. Numer. Anal.*, 38(3):837–875, 2000.
- [38] Steven E. Shreve. *Stochastic calculus for finance. II: Continuous-time models*. New York, NY: Springer, 2004.